Staying Alive

Connection Path Reselection at the Edge

Presenting: Raul Landa
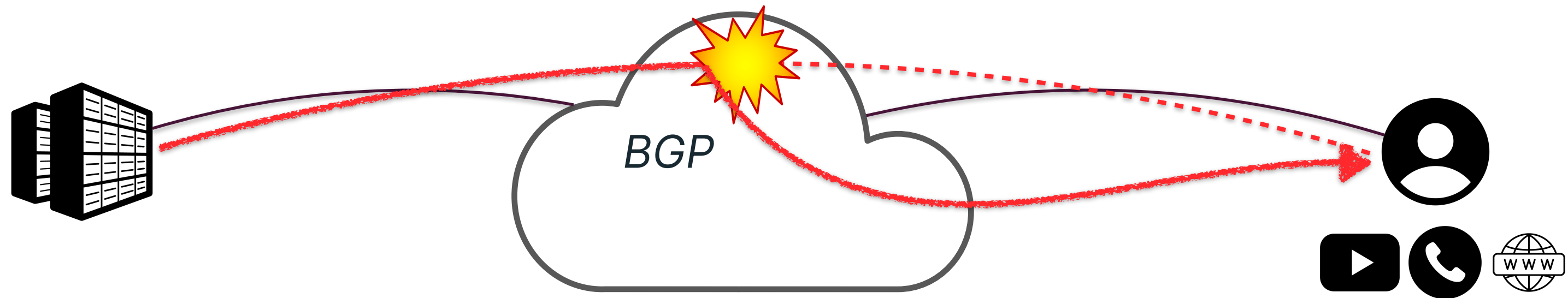
Raúl Landa
Lorenzo Saino
Lennert Buytenhek
João Taveira Araújo

March 2021

Path failure recovery

Fail, fail and fail again

‘Internet of things’ was mobilised for internet outage, says Dyn

Hacked devices linked to 'sophisticated' attack involving millions of IP addresses
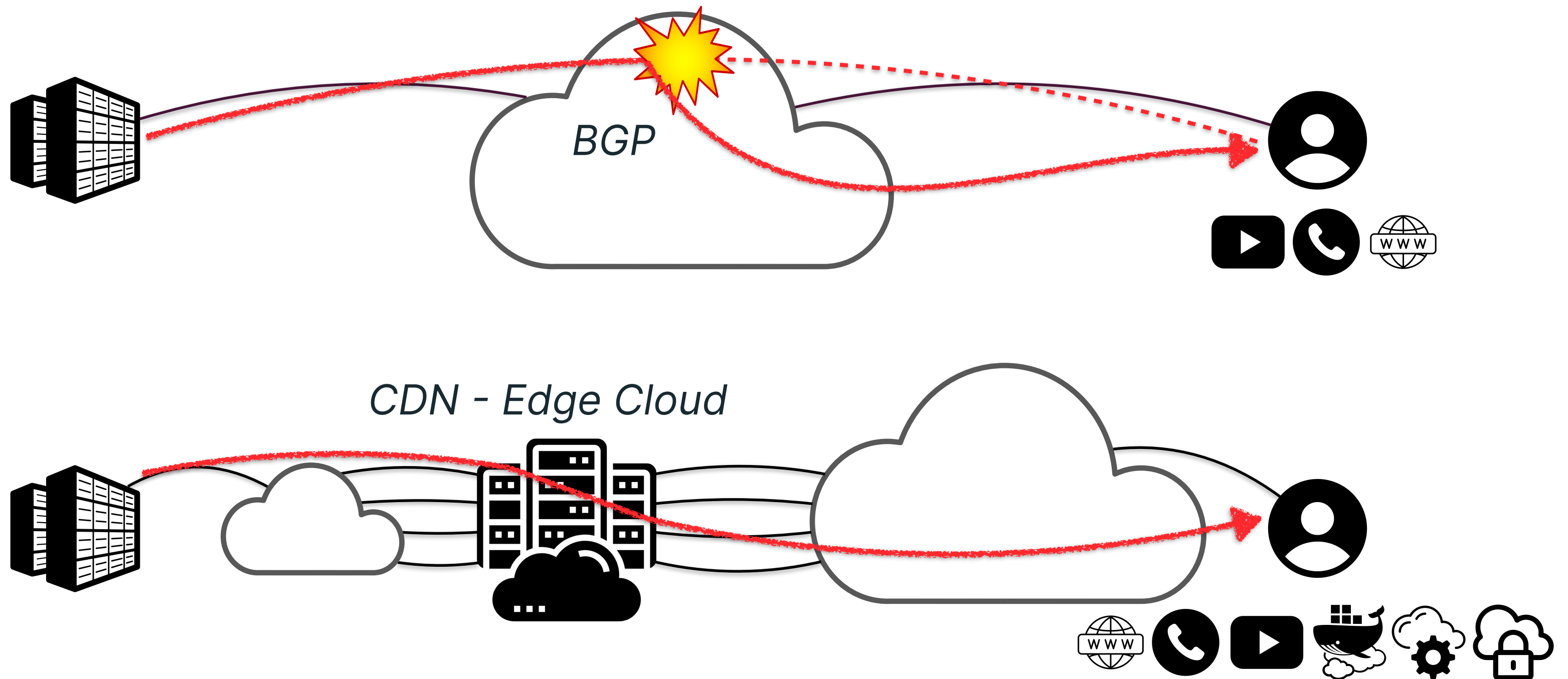
Financial Times, 2016

Internet connection cut between Europe, Asia and Africa

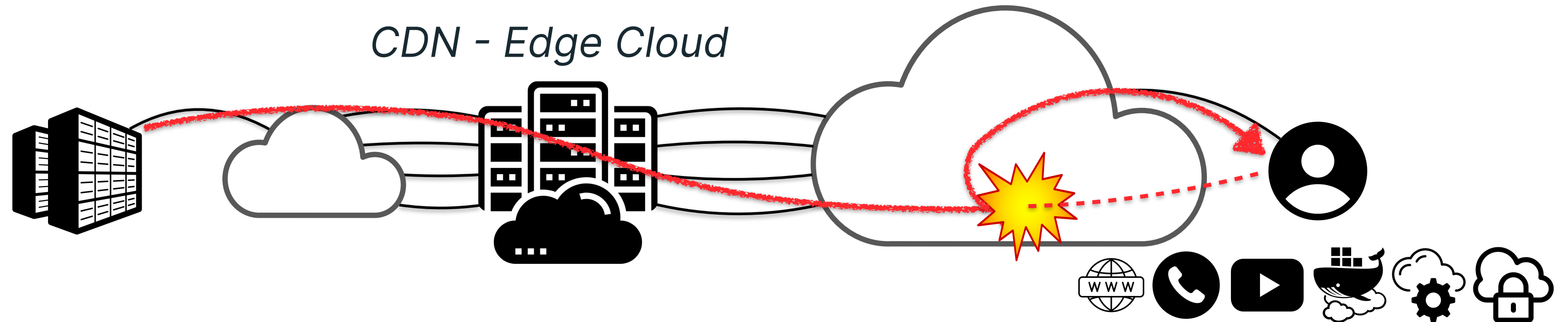'This outage is like a severed artery'

The Guardian, 2008

Path failure recovery

Enter CDN - Edge Cloud

BGP

CDN - Edge Cloud

Path failure recovery

Solutions that did not pan out

CDN - Edge Cloud

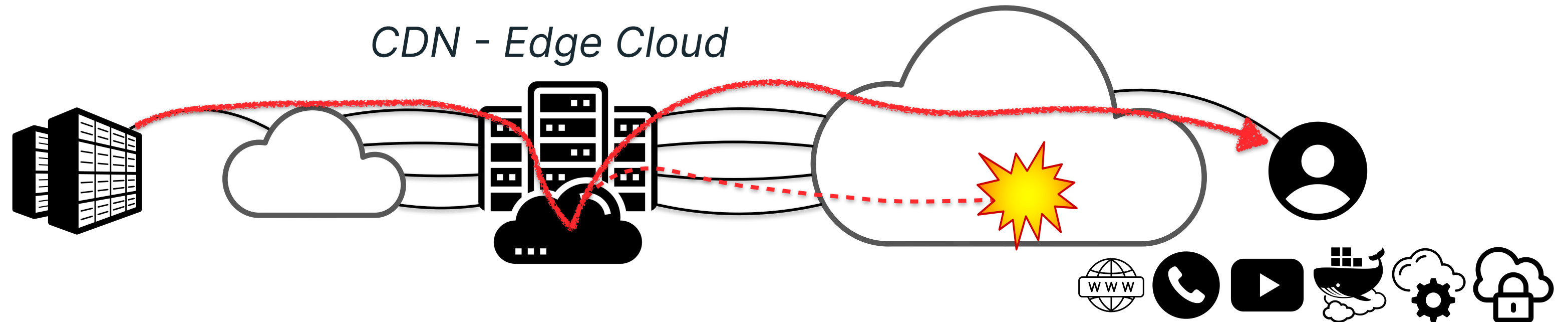*BGP*

- Slow interdomain convergence
- Can only detect a subset of failures

Path failure recovery

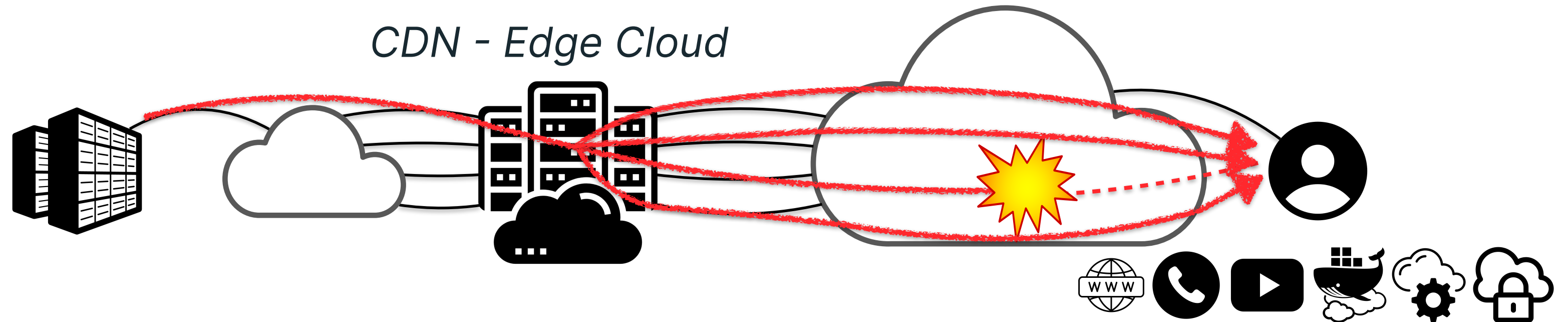Solutions that did not pan out

CDN - Edge Cloud

BGP

Intelligent Traffic Engineering

- Slow interdomain convergence
- Can only detect a subset of failures
- Most failures are sub-prefix and short lived
- Difficult to deploy using commodity switches

Path failure recovery

Solutions that did not pan out

CDN - Edge Cloud

BGP

Intelligent Traffic Engineering

Multipath transport

- Slow interdomain convergence
- Can only detect a subset of failures

- Most failures are sub-prefix and short lived
- Difficult to deploy using commodity switches

- Not widely deployed; requires client side support
- Only helps with already established connections
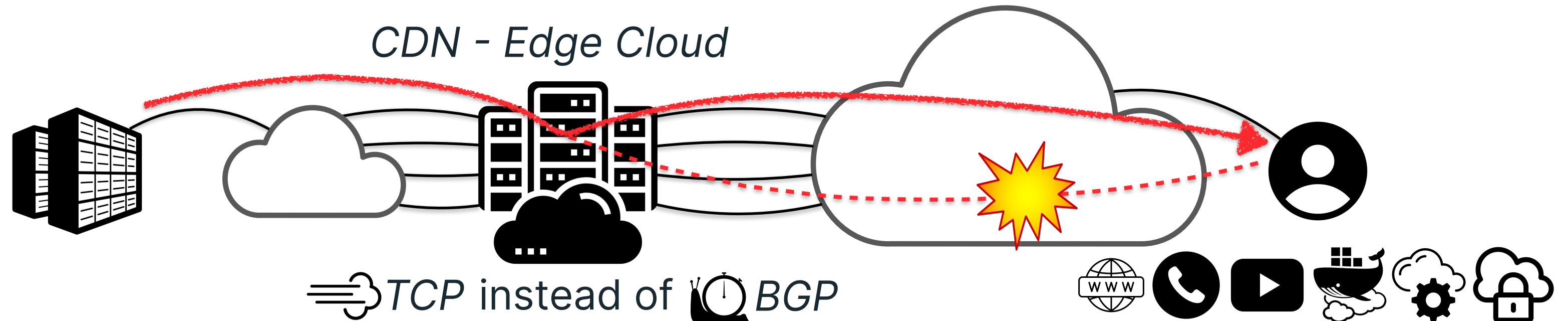
Path failure recovery

Re-examining the fundamentals
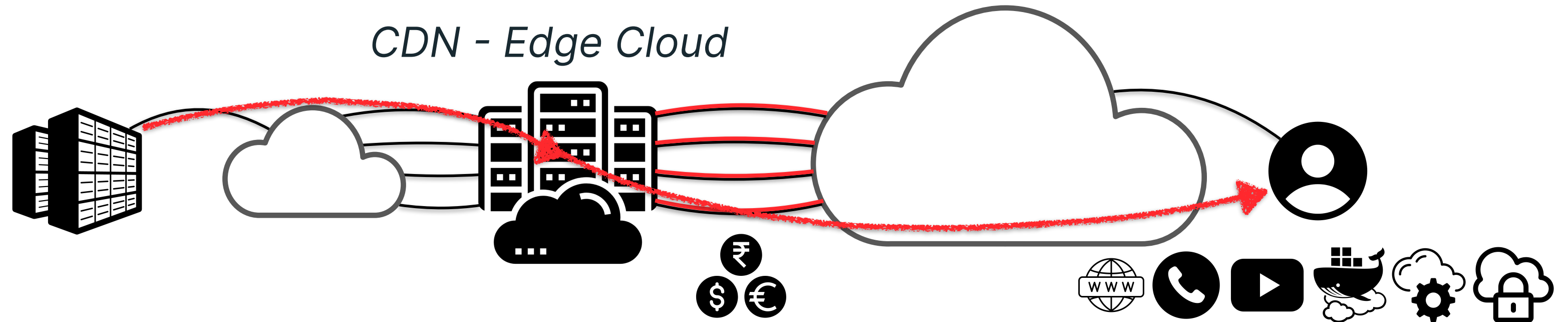
CDN - Edge Cloud

- Access to path diversity through multihoming
- End-to-end visibility of connection performance and state
- Economic incentives to address reliability

Path failure recovery

The idea behind Connection Path Reselection (CPR)

CDN - Edge Cloud

TCP instead of *BGP*

- Access to path diversity through multihoming
- End-to-end visibility of connection performance and state
- Economic incentives to address reliability
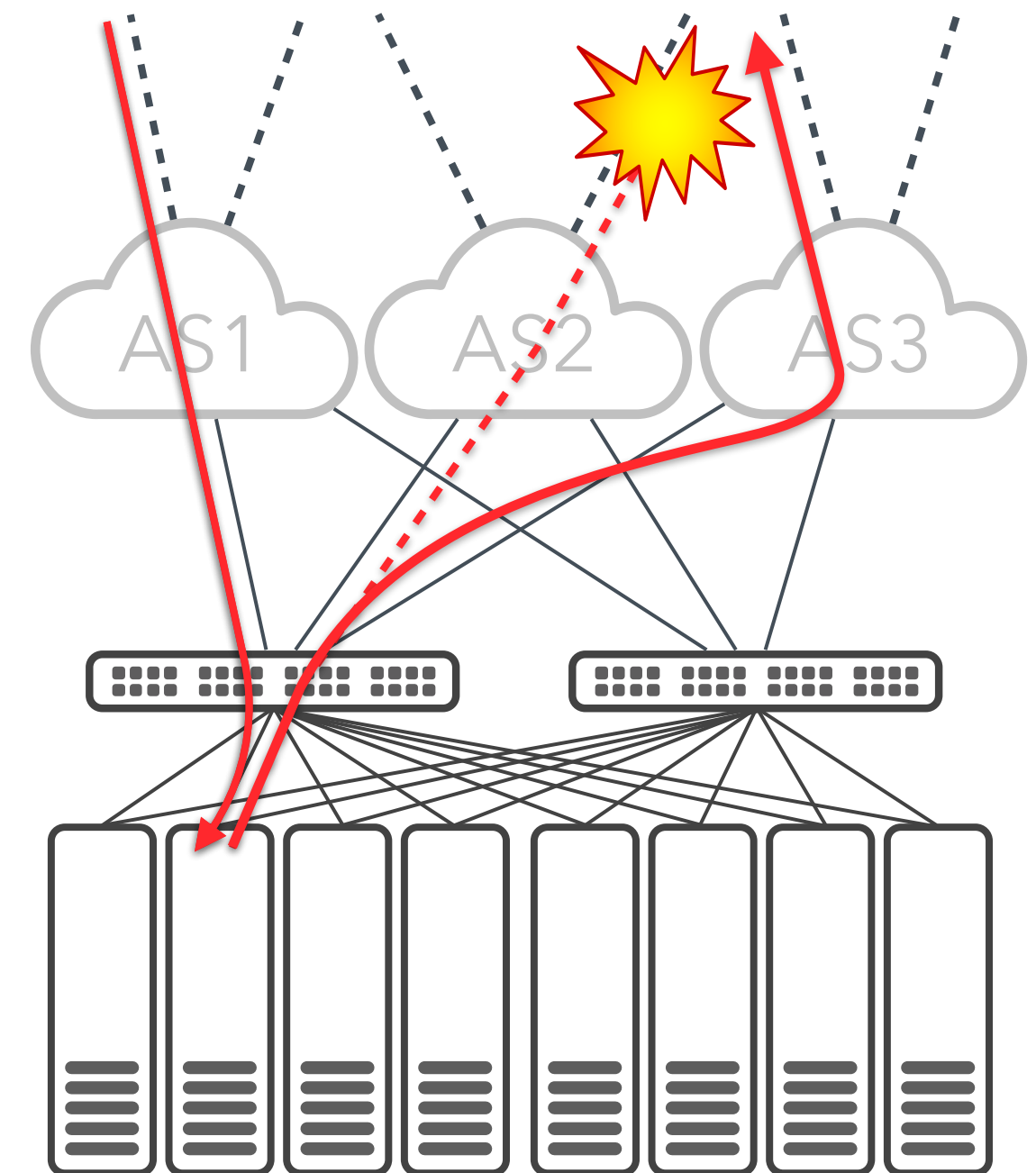- **Use transport layer information to drive egress routing**

What is CPR?

A high-level overview

A Linux kernel patch that

- *Monitors* TCP connections;
- *Detects* when the path associated with a connection has failed;

Impairment detection

- *Re-routes* outbound traffic for that connection; and

Path reselection

- *Measures* the aggregate effects of rerouting.

Impairment detection

Stalls and forward progress
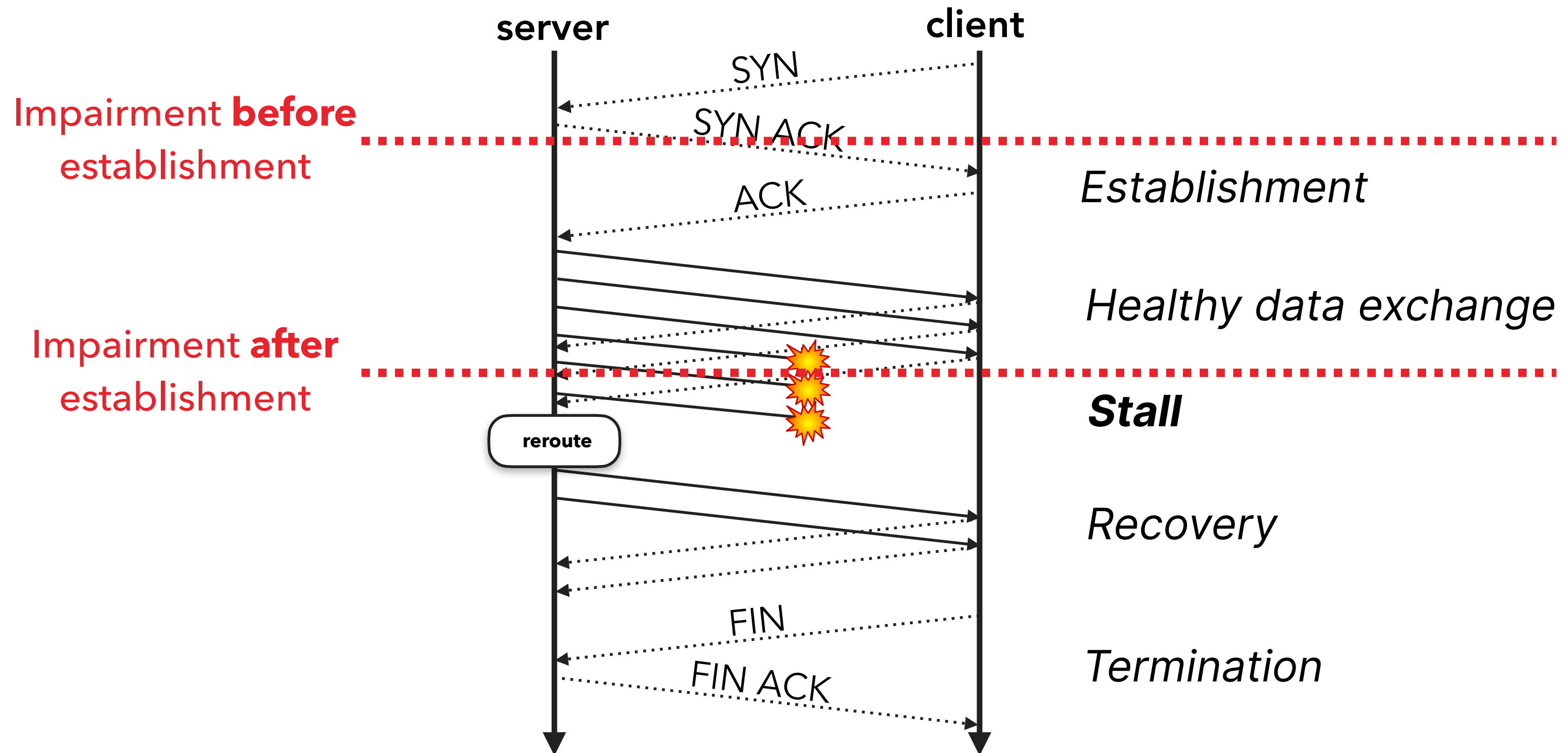
Impairment **before** establishment
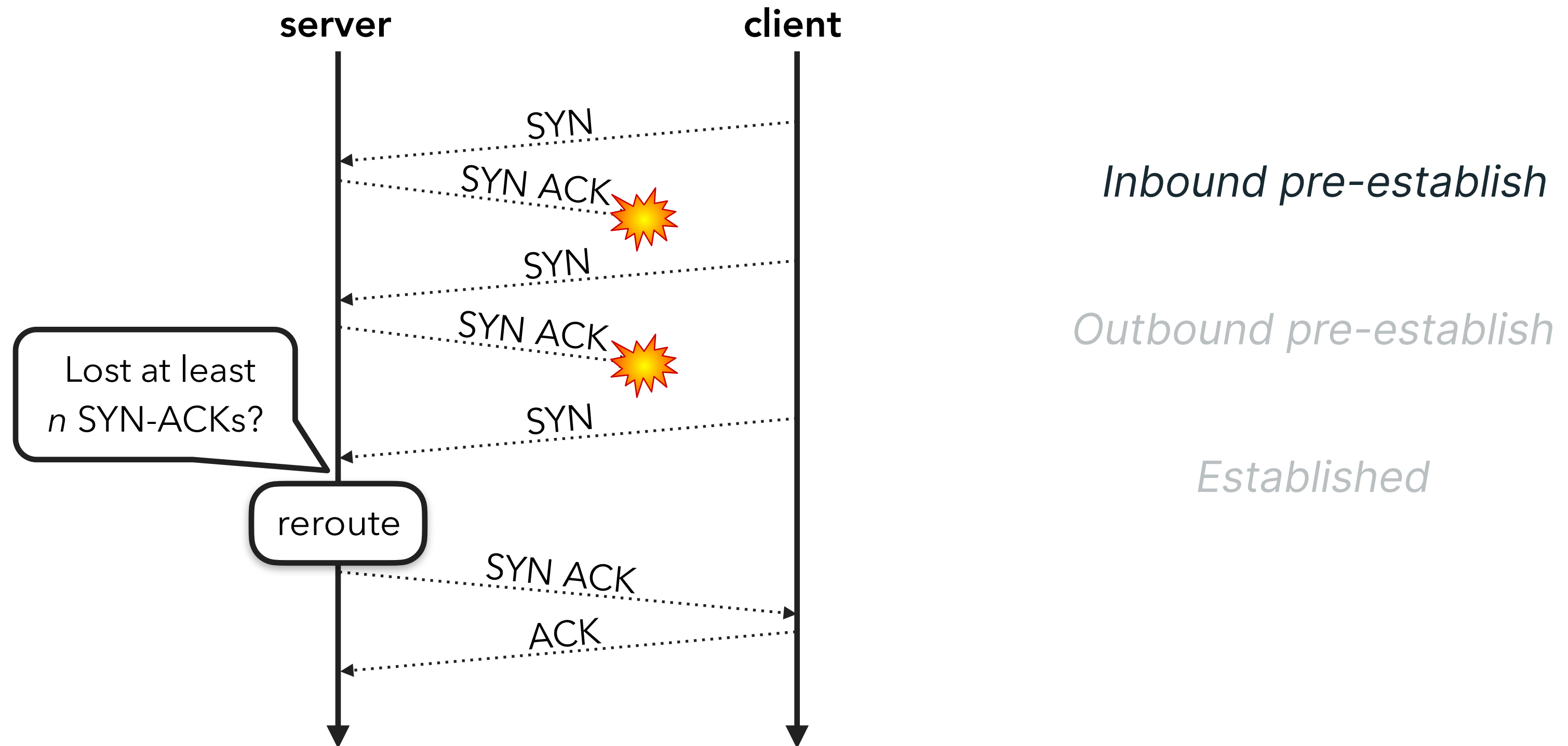
Impairment **after** establishment

reroute

server   client

SYN

SYN ACK

ACK

Establishment

Healthy data exchange

Stall

Recovery
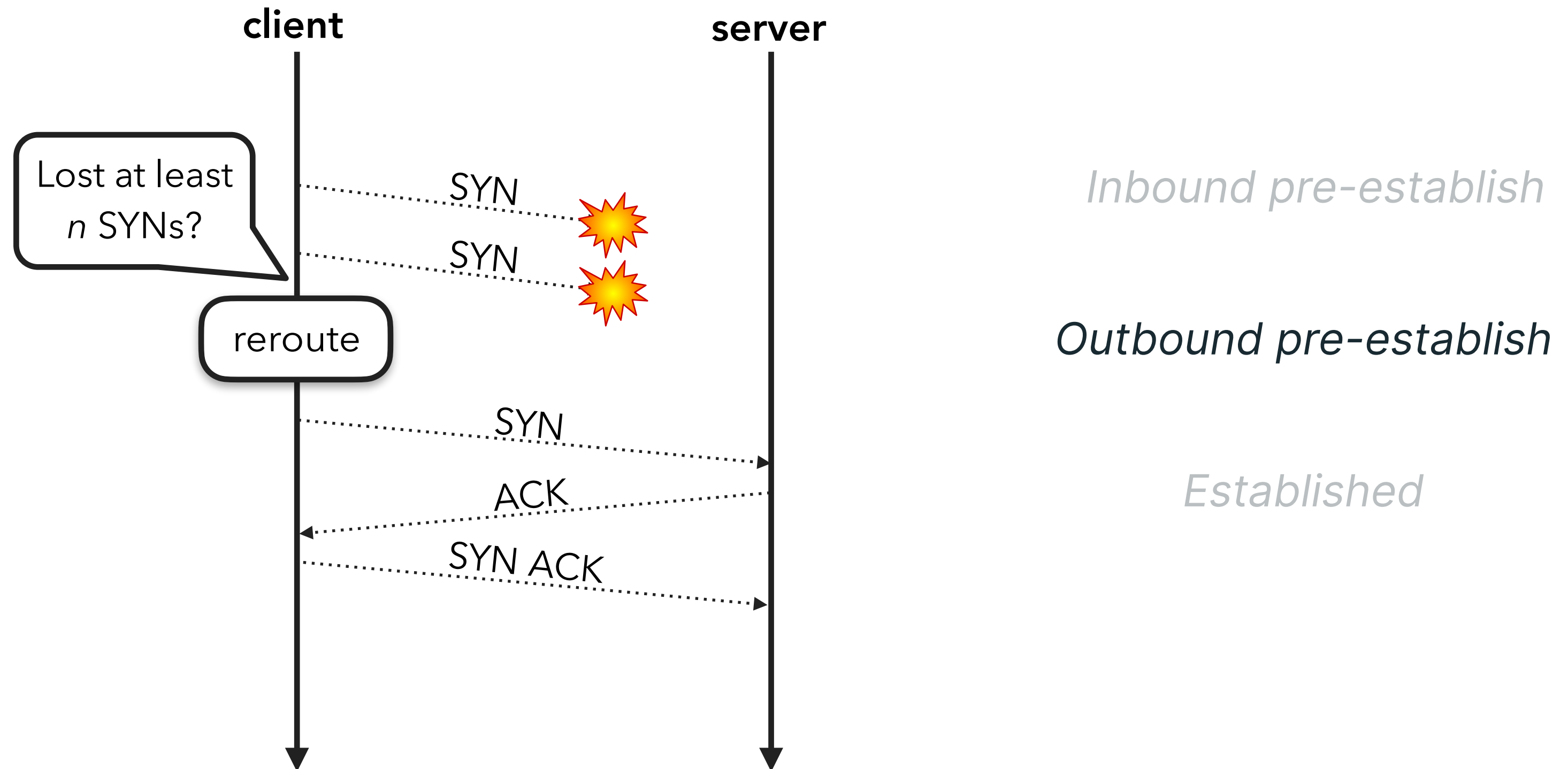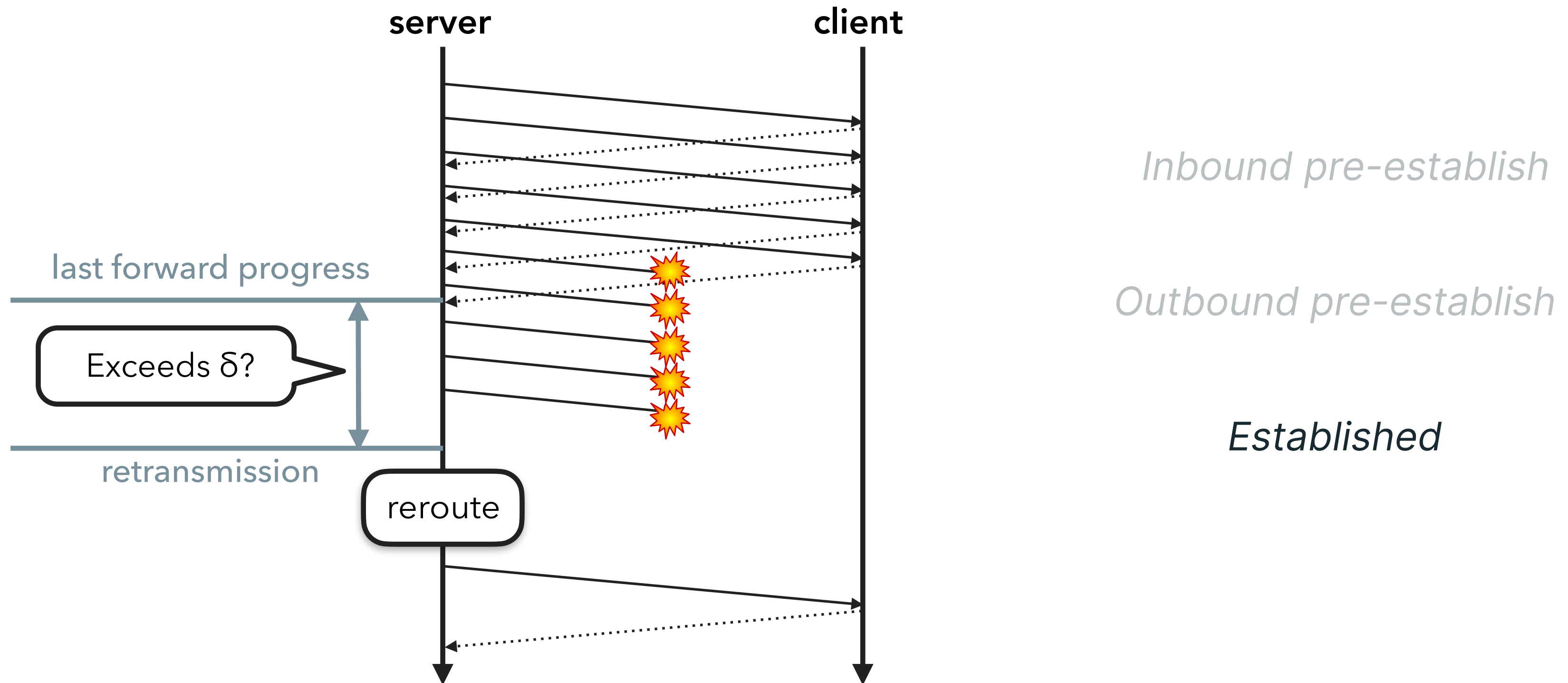
FIN

FIN ACK

Termination

Impairment detection

The mechanism used depends on connection lifecycle stage

Impairment detection

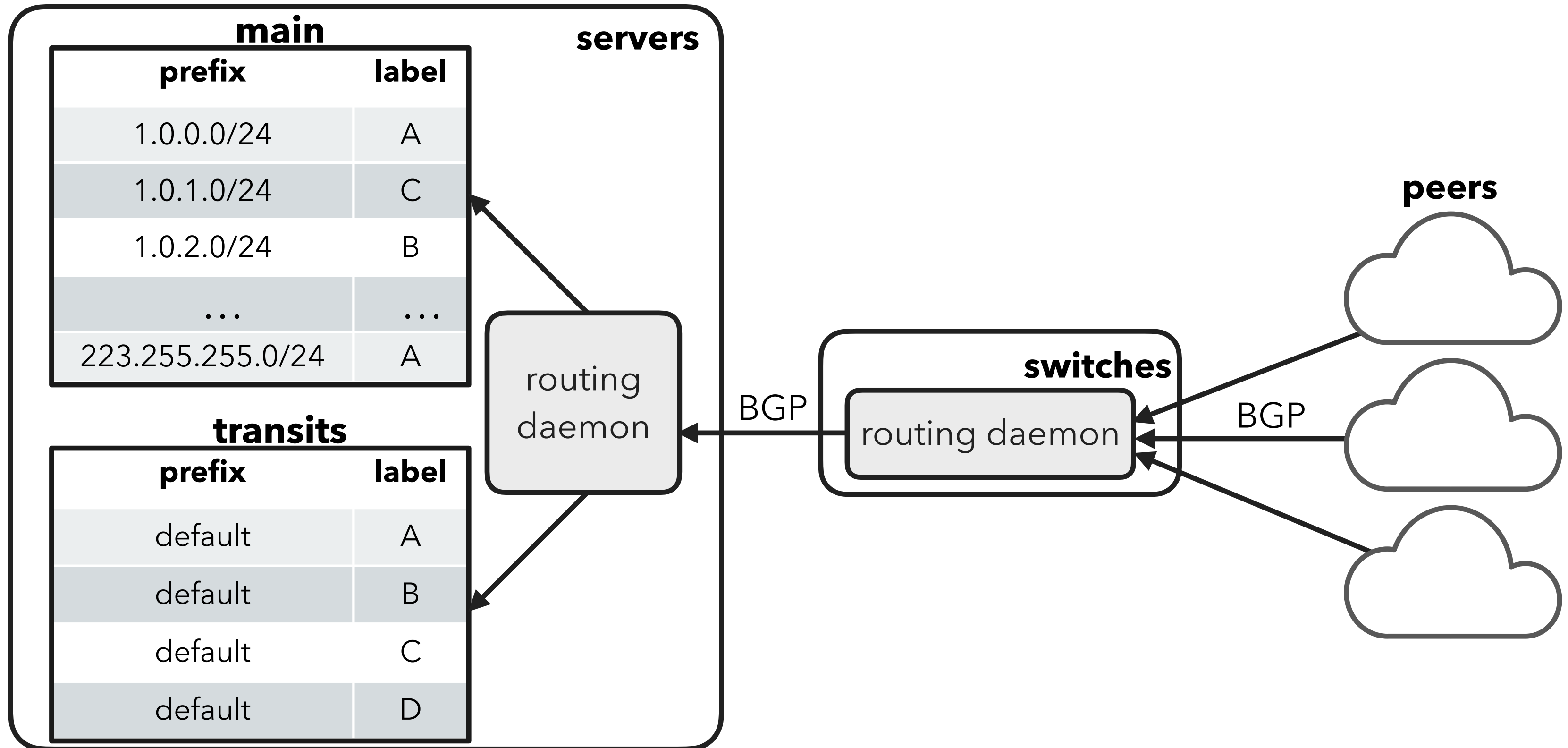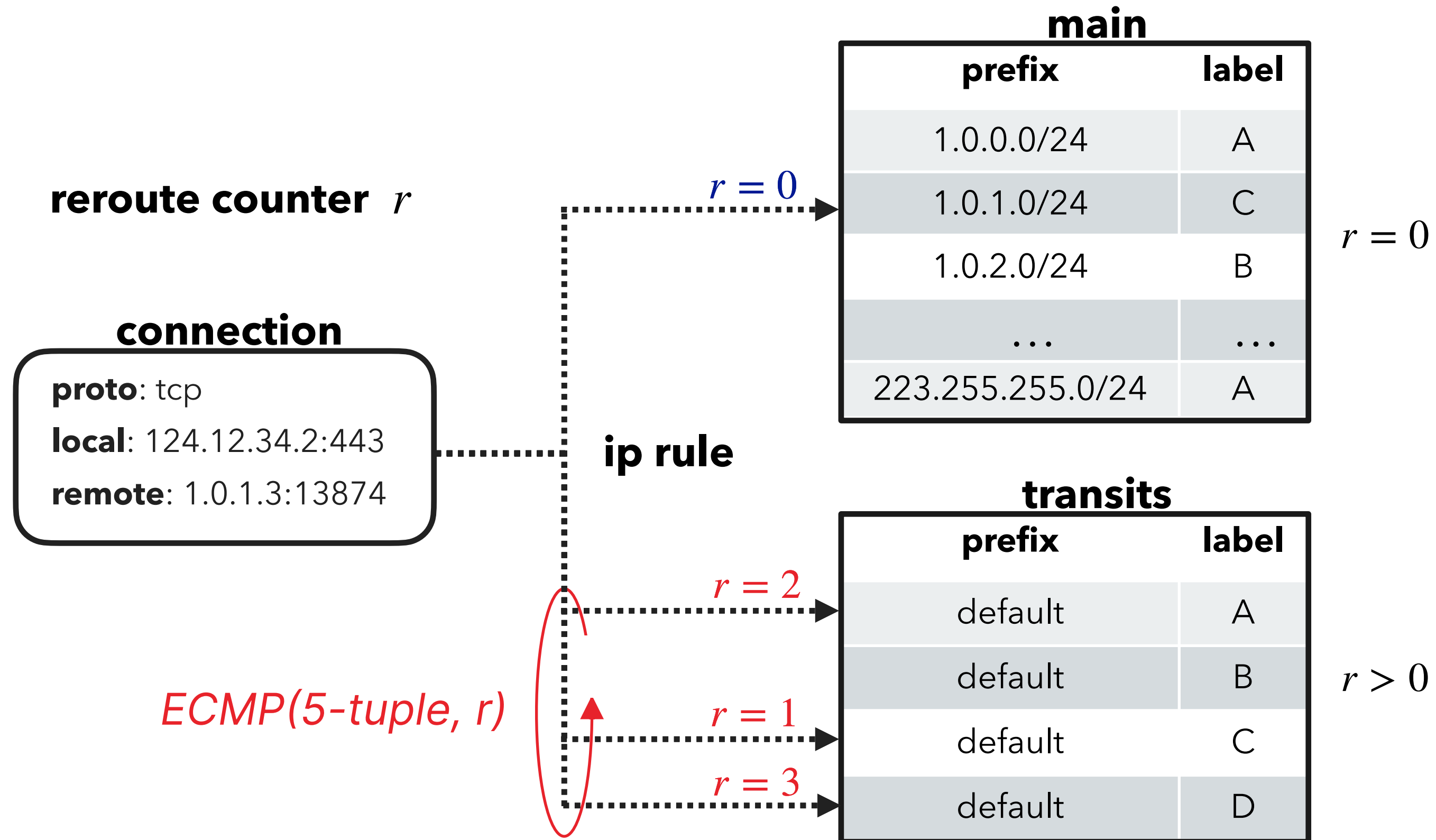The mechanism used depends on connection lifecycle stage

Impairment detection

The mechanism used depends on connection lifecycle stage

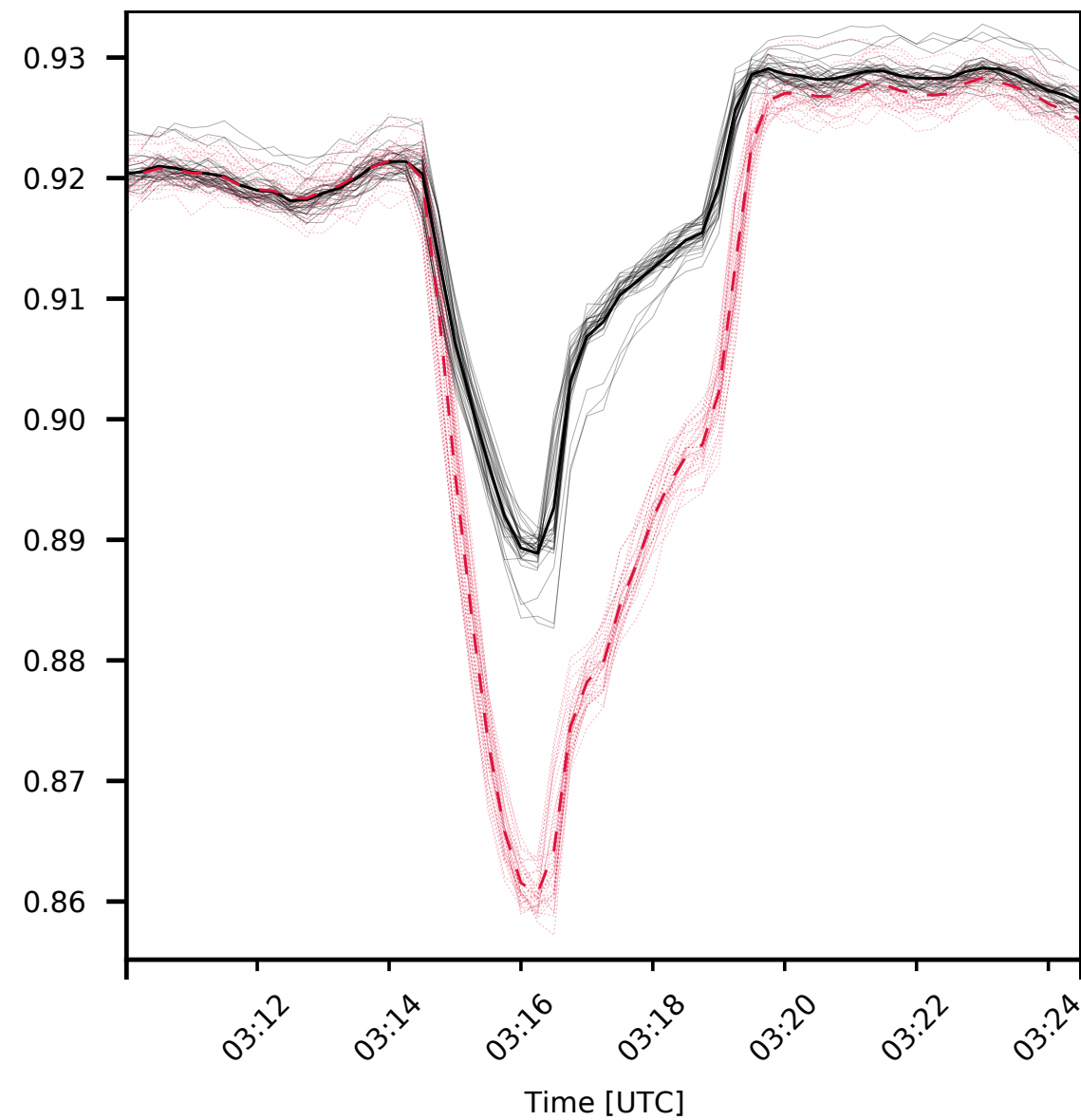# Path reselection

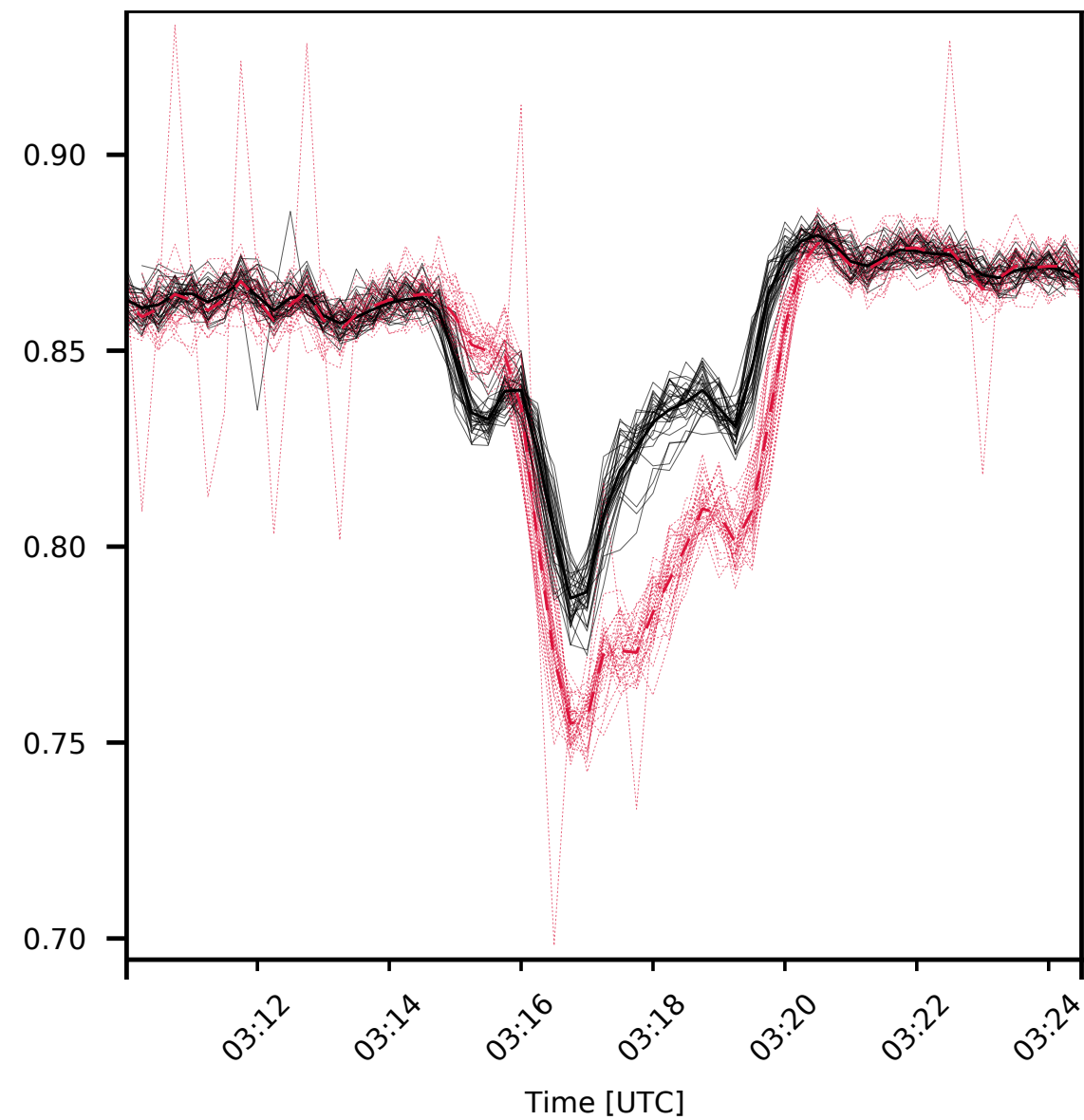Routing architecture

Path reselection

Rerouting

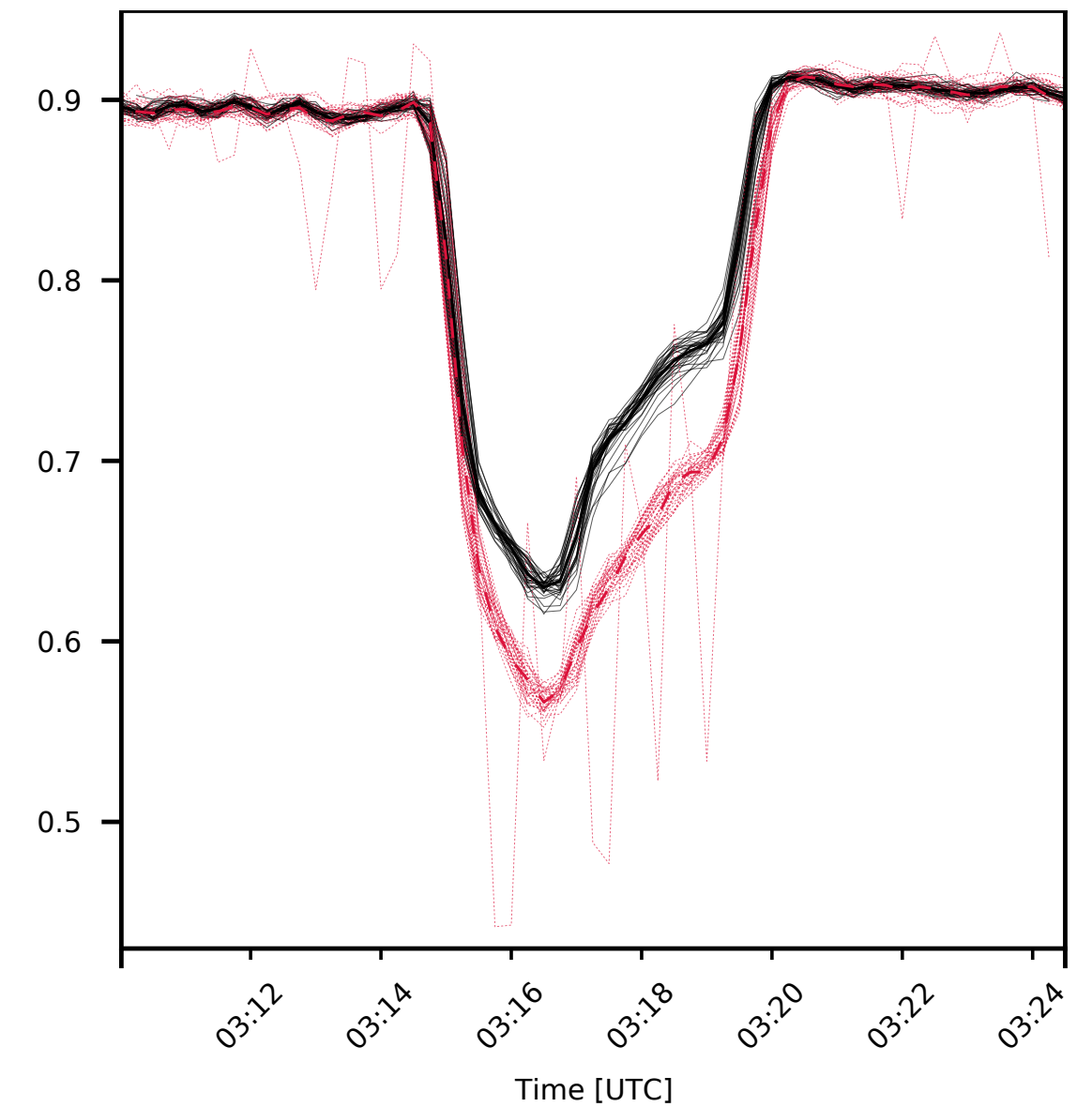# CPR in action

Measuring impact

● Treatment　　　● Control

$$\pi_\bullet(\text{HEALTHY})$$

$$p^o_\bullet(\text{HEALTHY}|\text{SYNRCVD})$$

$$p^o_\bullet(\text{CLOSE}|\text{HEALTHY})$$

In closing...

There's much more in the paper!

- How to tune the parameters for a production deployment

- Evaluating non-harm in the steady state

- Evaluating benefit during stall events, and tracking this benefit globally

- Path diversity from edge cloud PoPs

- Operational considerations, including Traffic Engineering and QUIC

- And many more contributions!

- For followup questions or comments please contact us at rlanda@fastly.com and lsaino@fastly.com

- And to the NSDI organizers, audience and every single person that helped to shape this paper...

Thank you!